CENTRE DE RECHERCHE
SUR LES TRANSPORTS
(C.R.T.)
CENTRE FOR RESEARCH
ON TRANSPORTATION

UNIVERSITÉ DE MONTRÉAL
ÉCOLE DES HAUTES ÉTUDES COMMERCIALES
ÉCOLE POLYTECHNIQUE

CENTRE DE RECHERCHE SUR LES TRANSPORTS
UNIVERSITÉ DE MONTRÉAL
C.P. 6128, SUCCURSALE A,
MONTRÉAL
CANADA H3C 3J7

TÉLÉPHONE: (514) 343-7575
TÉLEX: (21) 05-561325
EMAIL: crt@crt.umontreal.ca
TÉLÉCOPIEUR (FAX): (514) 343-7121

92-02300/

CRT-808

Expected Distances Between Two Random Points in Rectangles and Rectangular Parallelepipeds

by
Brahim Gaboune[1]
Gilbert Laporte[1,2]
François Soumis[1,2]

[1] GERAD, École des Hautes Études Commerciales de Montréal, 5255, avenue Decelles, Montréal, Canada, H3T 1V6

[2] Centre de recherche sur les transports, Université de Montréal, C.P. 6128, succursale A, Montréal (Québec) H3C 3J7

Centre de recherche sur les transports — Publication #808
GERAD — Cahier #G-91-52

January 1992

Dépôt légal – 1er trimestre 1992

Abstract

In this paper, the expected distance between two random points in a rectangle or in a rectangular parallelepiped is computed under three different metrics: the Manhattan metric, the Euclidean metric, and the Chebychev metric.

Keywords: Expected distances.

Résumé

Dans cet article, on calcule la distance moyenne entre deux points aléatoires dans un rectangle ou dans un parallélépipède rectangle sous trois métriques différentes: la métrique de Manhattan, la métrique euclidienne et la métrique de Chebychev.

Mots-clefs: Distances moyennes.

Introduction

The object of this paper is to provide analytical expressions for the expected distance between two randomly and uniformly distributed points in a rectangle of sides a and b, and in a rectangular parallelepiped of sides a, b and c, under three different metrics: the Manhattan metric (L_1), the Euclidean metric (L_2), and the Chebychev metric (L_∞). These metrics are defined as follows. For two points $P = (p_1, \ldots, p_r)$ and $Q = (q_1, \ldots, q_r)$ in r dimensions, the L_p distance between P and Q is

$$L_p = \left(\sum_{i=1}^{r} |p_i - q_i|^p \right)^{1/p},$$

where $p = 1$ yields the Manhattan metric, $p = 2$ yields the Euclidean metric, and $L_\infty = \sup_{i \in \{1,\ldots,r\}} (|p_i - q_i|)$ defines the Chebychev metric.

Average distance results in the plane are useful in a variety of distribution and production planning problems using geometric probabilities[1,2]. In particular, the L_1 and L_2 expected distances are often used in the context of vehicle routing, and in relation with the dispatching of emergency vehicles[2]. The L_∞ metric is more commonly employed in the determination of optimal paths in automatic punching or drilling operations. In the latter case, n holes must be drilled at specific locations on a rectangular metallic sheet by a press drill mounted on a fixed support. The sheet is moved n times to position the hole locations under the drill. For this, two independent motors operating simultaneously shift the sheet parallel to itself in each direction. The interpoint travel time is then equal to the maximum of two one-dimensional travel times.

In three dimensions, average distance problems occur mainly in manufacturing contexts. Consider again a drilling operation where, instead of a single press drill, there is a carousel that rotates to provide an appropriate drill. The problem of determining an optimal drilling sequence is then a three-dimensional travelling salesman problem. An example of such a system is provided in Chauny, Haurie, Wagneur and Loulou[3]. Depending on whether movements in each of the three dimensions are parallel or sequential, several operating systems exist. The distance (travel time) can be defined for a variety of metrics. More specifically, if X, Y and Z represent the distance between two points in each of three dimensions, then the overall distance L between these two points can be defined as $L = X + Y + Z$, $L = (X^2 + Y^2 + Z^2)^{1/2}$,

$L = \max(X, Y, Z)$, $L = \left[(X^2 + Y^2)^{1/2} + Z\right]$, $L = \max\left[(X^2 + Y^2)^{1/2}, Z\right]$, etc. Our analysis will be restricted to the first three cases which represent pure L_1, L_2 and L_∞ three-dimensional metrics.

In the case of the Manhattan metric in two or three dimensions, the average distance result is easily obtained from the additivity of expected values. It will be reported here for the sake of completeness. The average distance for the two dimensional Euclidean metric is more difficult to compute. An analytical expression is derived in Ghosh[4], but this result is not so well known and numerical approximations are generally used instead (see, e.g., Eilon, Watson-Gandy and Christofides[5], Larson and Odoni[2] and Hsu[6]). We provide in this paper a new but equivalent expression for this average distance. Results for the two-dimensional Chebychev metric, and three-dimensional average distance results under the L_2 and L_∞ metrics have never to our knowledge been reported before. We believe these can prove useful to researchers in the fields of distribution and of flexible manufacturing.

In addition to the above references, our work can be seen as complementary to that of Christofides and Eilon[7], Fairthorne[8], Murchland[9], Stone[10] and Vaughan[11] who have studied related but different expected distance problems.

Manhattan metric

The following notation will be used throughout the paper. Denote by (X_i, Y_i) or by (X_i, Y_i, Z_i) a point in a rectangle or in a rectangular parallelepiped. Consider two such points and define $X = |X_1 - X_2|$, $Y = |Y_1 - Y_2|$, $Z = |Z_1 - Z_2|$. Also let L denote the average distance between two uniformly randomly distributed points in the rectangle or in the parallelepiped.

For the Manhattan metric, the computation of $E(L)$ is relatively simple. First, for $0 \le x \le a$, the distribution function of X is given by

$$\begin{aligned} F_X(x) &= P\big(|X_1 - X_2| \le x\big) \\ &= 1 - \big[P(X_2 \ge X_1 + x) + P(X_2 \le X_1 - x)\big] \\ &= 1 - \left[\int_0^{a-x} \int_{x_1+x}^{a} f(x_1, x_2)\, dx_2\, dx_1 + \int_x^a \int_0^{x_1-x} f(x_1, x_2)\, dx_2\, dx_1\right], \end{aligned}$$

where $f(x_1, x_2)$, the joint probability function of X_1 and X_2, is defined by

$$f(x_1, x_2) = \begin{cases} 1/a^2 & \text{if} \quad 0 \le x_1 \le a \text{ and } 0 \le x_2 \le a \\ 0 & \text{otherwise,} \end{cases}$$

since X_1 and X_2 are independent. Therefore,

$$F_X(x) = \begin{cases} 0 & \text{if} \quad x \le 0 \\ 1 - (x-a)^2/a^2 & \text{if} \quad 0 \le x \le a \\ 1 & \text{if} \quad x \ge a, \end{cases}$$

so that the density function of X is triangular over $[0, a]$:

$$f_X(x) = \begin{cases} \frac{2}{a}\left(1 - \frac{x}{a}\right) & \text{if} \quad 0 \le x \le a \\ 0 & \text{otherwise .} \end{cases} \tag{1}$$

Similar results hold for Y and Z. The expected values of these variables are then easily computed as $E(X) = a/3$, $E(Y) = b/3$ and $E(Z) = c/3$. It follows that $E(L) = E(X+Y) = (a+b)/3$ in two dimensions, and $E(L) = E(X+Y+Z) = (a+b+c)/3$ in three dimensions.

Two-dimensional Euclidean metric

Using (1) and its counterpart for Y, $E(L) = E\left[(X^2+Y^2)^{1/2}\right]$ can be written as

$$E(L) = \frac{4}{ab}\int_0^b \int_0^a \sqrt{x^2+y^2}\left(1 - \frac{x}{a}\right)\left(1 - \frac{y}{b}\right) dx\, dy.$$

The integration domain is decomposed into two triangles, as shown in Figure 1, and the value of $E(L)$ is the sum of two integrals I and J, each defined on one triangle. We first compute I:

$$\begin{aligned} I &= \frac{4}{ab}\int_0^a \int_0^{\frac{b}{a}x} \sqrt{x^2+y^2}\left(1 - \frac{x}{a}\right)\left(1 - \frac{y}{b}\right) dy\, dx \\ &= \frac{2}{ab}\int_0^a \left(1 - \frac{x}{a}\right) G(x)\, dx, \end{aligned}$$

where

$$
\begin{aligned}
G(x) &= \int_0^{\frac{b}{a}x} \left(2\sqrt{x^2+y^2} - \frac{2}{b}y\sqrt{x^2+y^2}\right) dy \\
&= \left[y\sqrt{x^2+y^2} + x^2 \ln\left(y+\sqrt{x^2+y^2}\right) - \frac{2}{3b}(x^2+y^2)^{3/2}\right]_{y=0}^{y=\frac{b}{a}x} \\
&= \left[\frac{b\sqrt{a^2+b^2}}{a^2} + \ln\left(\frac{b+\sqrt{a^2+b^2}}{a}\right)\right] x^2 - \frac{2}{3b}\left(\frac{(a^2+b^2)^{3/2}}{a^3} - 1\right) x^3.
\end{aligned}
$$

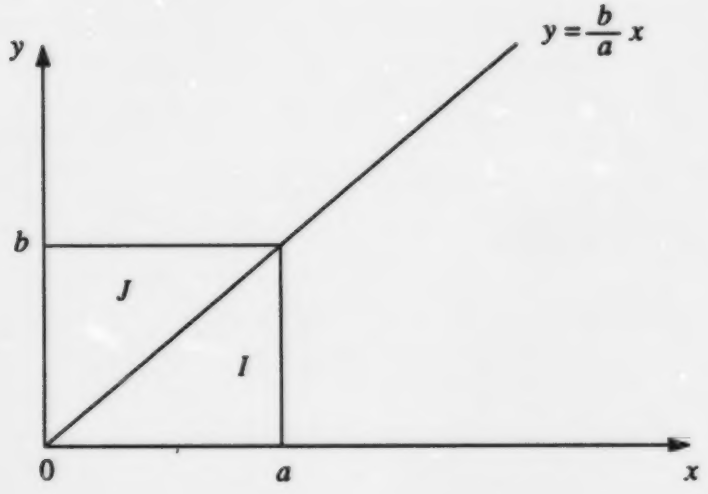

Figure 1. Decomposition of the integration domain into two triangles

Therefore,

$$I = \frac{2}{ab}\int_0^a \left(1 - \frac{x}{a}\right)\left(Ax^2 - \frac{2}{3b}Bx^3\right) dx, \tag{2}$$

where

$$A = \frac{b\sqrt{a^2+b^2}}{a^2} + \ln\left(\frac{b+\sqrt{a^2+b^2}}{a}\right)$$

and

$$B = \frac{(a^2+b^2)^{3/2}}{a^3} - 1.$$

Letting $t = \dfrac{x}{a}$ in (2), we have

$$\begin{aligned} I &= \frac{2a^2}{b}\int_0^1 (1-t)\left(At^2 - \frac{2a}{3b}Bt^3\right)dt \\ &= \frac{a^2}{b}\left(\frac{A}{6} - \frac{a}{15b}B\right) \\ &= \frac{1}{6}\sqrt{a^2+b^2} + \frac{a^2}{6b}\ln\left(\frac{b+\sqrt{a^2+b^2}}{a}\right) - \frac{(a^2+b^2)^{3/2}}{15b^2} + \frac{a^3}{15b^2}. \end{aligned} \tag{3}$$

By permuting a and b in (3), we find

$$J = \frac{1}{6}\sqrt{a^2+b^2} + \frac{b^2}{6a}\ln\left(\frac{a+\sqrt{a^2+b^2}}{b}\right) - \frac{(a^2+b^2)^{3/2}}{15a^2} + \frac{b^3}{15a^2}.$$

Therefore,

$$\begin{aligned} E(L) = I + J &= \frac{1}{3}\sqrt{a^2+b^2} + \frac{a^2}{6b}\ln\left(\frac{b+\sqrt{a^2+b^2}}{a}\right) \\ &\quad + \frac{b^2}{6a}\ln\left(\frac{a+\sqrt{a^2+b^2}}{b}\right) - \frac{(a^2+b^2)^{5/2}}{15a^2b^2} + \frac{a^5+b^5}{15a^2b^2}, \end{aligned}$$

which reduces to

$$E(L) = \frac{d}{15}\left\{2 + \sqrt{2} + 5\ln(1+\sqrt{2})\right\} = 0.5214d$$

for a square of side d.

Two-dimensional Chebychev metric

To compute $E(L) = E\big[\max(X,Y)\big]$, we again require the density functions of X and Y (see (1)). Assuming $a \le b$, the distribution function of L is:

$$F_L(\ell) = P(L \le \ell) = P(X \le \ell)P(Y \le \ell),$$

since X and Y are independent. Therefore,

$$F_L(\ell) = \begin{cases} 0 & \text{if} \quad \ell \le 0 \\ \displaystyle\int_0^\ell \frac{2(a-x)}{a^2}\,dx \int_0^\ell \frac{2(b-y)}{b^2}\,dy & \text{if} \quad 0 \le \ell \le a \\ \displaystyle\int_0^\ell \frac{2(b-y)}{b^2}\,dy & \text{if} \quad a \le \ell \le b \\ 1 & \text{if} \quad \ell \ge b. \end{cases}$$

The density function of $L = \max(X, Y)$ can be written as

$$f_L(\ell) = \begin{cases} \dfrac{\ell}{a^2 b^2}\left[4\ell^2 - 6(a+b)\ell + 8ab\right] & \text{if} \quad 0 \le \ell \le a \\ \dfrac{2(b-\ell)}{b^2} & \text{if} \quad a \le \ell \le b \\ 0 & \text{otherwise.} \end{cases}$$

Thus, $E(L) = \displaystyle\int_0^b \ell f_L(\ell)\,d\ell = (-a^3 + 5a^2 b + 10b^3)/30b^2$, which reduces to

$$E(L) = 7d/15$$

for a square of side d.

Three-dimensional Euclidean metric

We now consider a rectangular parallelepiped of sides a, b and c with $a \le b \le c$. Using (1) again, we have

$$\begin{aligned} E(L) &= \frac{8}{abc}\int_0^c \int_0^b \int_0^a \sqrt{x^2+y^2+z^2}\left(1-\frac{x}{a}\right)\left(1-\frac{y}{b}\right)\left(1-\frac{z}{c}\right)\,dx\,dy\,dz \\ &= 8\int_0^1 \int_0^1 \int_0^1 \sqrt{a^2x^2+b^2y^2+c^2z^2}(1-x)(1-y)(1-z)\,dx\,dy\,dz. \end{aligned}$$

We know no closed form solution for this integral, but it can be computed numerically for some values of the parameters. For example, using the MathCAD[12] package, we obtain

$$E(L) = 8(0.0827134705)\, d = 0.6617\, d$$

for a cube of side d.

Three-dimensional Chebychev metric

Finally, we compute $E(L) = E\big[\max(X, Y, Z)\big]$ in a parallelepiped of sides a, b and c, with $a \le b \le c$. The distribution function of L is

$$F_L(\ell) = P(L \le \ell) = P(X \le x)P(Y \le y)P(Z \le z).$$

Again, the three independent variables X, Y and Z have similar triangular distributions (see (1)), so that

$$F_L(\ell) = \begin{cases} 0 & \text{if} \quad \ell \le 0 \\ \displaystyle\int_0^\ell \frac{2}{a}\left(1-\frac{x}{a}\right) dx \int_0^\ell \frac{2}{b}\left(1-\frac{y}{b}\right) dy \int_0^\ell \frac{2}{c}\left(1-\frac{z}{c}\right) dz & \text{if} \quad 0 \le \ell \le a \\ \displaystyle\int_0^\ell \frac{2}{b}\left(1-\frac{y}{b}\right) dy \int_0^\ell \frac{2}{c}\left(1-\frac{z}{c}\right) dz & \text{if} \quad a \le \ell \le b \\ \displaystyle\int_0^\ell \frac{2}{c}\left(1-\frac{z}{c}\right) dz & \text{if} \quad b \le \ell \le c \\ 1 & \text{if} \quad \ell \ge c, \end{cases}$$

which reduces to

$$F_L(\ell) = \begin{cases} 0 & \text{if} \quad \ell \le 0 \\ \dfrac{1}{a^2b^2c^2}\left[8abc\,\ell^3 - 4\ell^4(ab+ac+bc) + 2\ell^5(a+b+c) - \ell^6\right] & \text{if} \quad 0 \le \ell \le a \\ \dfrac{1}{b^2c^2}\left[4bc\,\ell^2 - \ell^3(2c+2b) + \ell^4\right] & \text{if} \quad a \le \ell \le b \\ \dfrac{1}{c^2}\left(2c\ell - \ell^2\right) & \text{if} \quad b \le \ell \le c \\ 1 & \text{if} \quad \ell \ge c. \end{cases}$$

The density function of L can then be expressed as

$$f_L(\ell) = \begin{cases} \frac{1}{a^2b^2c^2}\left[24abc\,\ell^2 - 16\ell^3(ab+ac+bc) + 10\ell^4(a+b+c) - 6\ell^5\right] & \text{if} \quad 0 \le \ell \le a \\ \frac{\ell}{b^2c^2}\left[4\ell^2 - 6(b+c)\ell + 8bc\right] & \text{if} \quad a \le \ell \le b \\ \frac{2(c-\ell)}{c^2} & \text{if} \quad b \le \ell \le c \\ 0 & \text{otherwise.} \end{cases}$$

Hence,

$$\begin{aligned} E(L) &= \int_0^c \ell f_L(\ell)\, d\ell \\ &= \frac{a^3}{b^2c^2}\left[6bc - \frac{16}{5}(ab+ac+bc) + \frac{5}{3}(a^2+ab+ac) - \frac{6}{7}a^2\right] \\ &\quad + \frac{1}{a^2c^2}\left[\frac{4}{5}(b^5-a^5) - \frac{3}{2}(b+c)(b^4-a^4) + \frac{8}{3}bc(b^3-a^3)\right] \\ &\quad + \frac{2}{c^2}\left[\frac{c}{2}(c^2-b^2) - \frac{1}{3}(c^3-b^3)\right], \end{aligned}$$

which reduces to

$$E(L) = 19d/35$$

for a cube of side d.

Conclusion

We have computed the average distance between two randomly and uniformly distributed points in a rectangle or in a rectangular parallelepiped under the L_1, L_2 and L_∞ metrics. These results are useful in a number of distribution and manufacturing contexts. The L_1 average distances are easily derived from the sum of one-dimensional expected interpoint values. The L_2 result for two dimensions has been presented before, but is not well known and several authors still use approximations instead of the closed from expression. In three dimensions, the L_2 expected distance cannot be easily expressed in closed form, but numerical approximations can be worked out for some values of the parameters. The L_∞ results have not apparently been reported before. These are particularly useful in the evaluation of drilling sequences in a number of flexible manufacturing settings.

Acknowledgement

This research was partially supported by Quebec government grant FCAR 91ER0290. Thanks are due to Michel Moran for his helpful advice.

References

1. C.F. Daganzo (1991) *Logistics Systems Analysis. Lecture Notes in Economics and Mathematical Systems* 361, M. Beckmann and W. Krelle (Eds.), Springer-Verlag, Berlin.

2. R.C. Larson and A.R. Odoni (1981) *Urban Operations Research*, Prentice-Hall, Englewood Cliffs, NJ.

3. F. Chauny, A. Haurie, E. Wagneur and R. Loulou (1987) Sequencing Punch Operations in a Flexible Manufacturing Cell: a Three-Dimensional Space Filling Curve Approach, *INFOR* 25, 26–45.

4. B. Ghosh (1951) Random Distances within a Rectangle and between Two Rectangles, *Bulletin of the Calcutta Mathematical Society* 43, 17–24.

5. S. Eilon, C.D.T. Watson-Gandy and N. Christofides (1971) *Distribution Management: Mathematical Modelling and Practical Analysis*, Griffin, London.

6. A.C. Hsu (1990) *Expected Distance between Two Random Points in a Polygon*, M. Sc. Dissertation, Department of Civil Engineering, Massachusetts Institute of Technology.

7. N. Christofides and S. Eilon (1969) Expected Distances in Distribution Problems, *Operational Research Quarterly* 20, 437–443.

8. D. Fairthorne (1965) Distances between Pairs of Points in Towns of Simple Geometrical Shapes, *Proceedings of the Second International Symposium on the Theory of Road Traffic Flow*, OECD, Paris, 391–406.

9. J.D. Murchland (1976) Solution to Problem 75–12, *SIAM Review* 18, 498.

10. R.E. Stone (1991) Some Average Distance Results, *Transportation Science* 25, 83–91.

11. R. Vaughan (1984) Approximate Formulas for the Average Distances Associated with Zones, *Transportation Science* 18, 231–244.

12. MathCAD Version 2.03 (1987), Mathsoft Inc.